



Q67867 Sequence2.ST25.txt  
SEQUENCE LISTING

<110> VAUGHAN, Paul R.  
GALANIS, Maria  
RAMSHAW, John A.M.  
WERKMEISTER, Jerome A.

<120> METHOD FOR PRODUCING, IN YEAST, A HYDROXYLATED TRIPLE HELICAL PROTEIN, AND  
YEAST HOST CELLS USEFUL IN SAID METHOD

<130> Q67867

<140> 10/023,831

<141> 2001-12-21

<150> US 09/297,269

<151> 1999-04-28

<150> P03310

<151> 1996-10-29

<150> P04306

<151> 1996-12-19

<150> PCT/AU97/00721

<151> 1997-10-29

<160> 52

<170> PatentIn version 3.1

<210> 1

<211> 26

<212> DNA

<213> Unknown

<220>

<223> Synthetic Construct

<400> 1

ctgtagagga tccccgggta cggagc

26

<210> 2

<211> 23

<212> DNA

<213> Unknown

<220>

<223> Synthetic Construct

<400> 2

ttatattgaa ttctcaaaaa ttc

23

<210> 3

<211> 32

<212> DNA

<213> Unknown

<220>

<223> Synthetic Construct

<400> 3

tgtaaaatta aaggatccca aagatgtggt at

32

<210> 4

<211> 30

<212> DNA

<213> Unknown

<220>

Q67867 Sequence2.ST25.txt

<223> Synthetic Construct

<400> 4  
gccgggatcc tgtcattcca atgacaacgt

30

<210> 5

<211> 44

<212> DNA

<213> Unknown

<220>

<223> Synthetic Construct

<400> 5  
actggacgga tcccgagcgc cccgcctgct ccgtgtccga catg

44

<210> 6

<211> 24

<212> DNA

<213> Unknown

<220>

<223> Synthetic Construct

<400> 6  
ggttctcctt ggtgacctcc cctt

24

<210> 7

<211> 24

<212> DNA

<213> Unknown

<220>

<223> Synthetic Construct

<400> 7  
gaaggggagg tcaccaagga gaac

24

<210> 8

<211> 36

<212> DNA

<213> Unknown

<220>

<223> Synthetic Construct

<400> 8

ccttcaggat cctattagac ttcattcttc aacagc

36

<210> 9

<211> 44

<212> DNA

<213> Unknown

<220>

<223> Synthetic Construct

<400> 9

actggacgga tcccgagcgc cccgcctgct ccgtctccga catg

44

<210> 10

<211> 35

<212> DNA

<213> Unknown

<220>

<223> Synthetic Construct

<400> 10

ccttcaggat cctattagac ttcattcttc acagc

35

<210> 11

<211> 39

<212> DNA

<213> Unknown

<220>

Q67867 Sequence2.ST25.txt

<223> Synthetic Construct

<400> 11

ctagttgaat tctacacaat gctgcccgc gctctgctg

39

<210> 12

<211> 36

<212> DNA

<213> Unknown

<220>

<223> Synthetic Construct

<400> 12

gcaatggaat tcttattaca gttcgtgcac agcttt

36

<210> 13

<211> 4

<212> PRT

<213> Yeast

<400> 13

His Asp Glu Leu  
1

<210> 14

<211> 5

<212> PRT

<213> Mammalian

<400> 14

Lys Asp Ala Glu Leu  
1 5

<210> 15

<211> 42

<212> DNA

<213> Unknown

<220>

<223> Collagen Type III

<400> 15

gctagcaagc ttggagctcc aggccactt gggattgctg gg

42

<210> 16

<211> 45

<212> DNA

<213> Unknown

<220>

<223> Collagen Type III

<400> 16

tcgcatctta gattataaaa agcaaacagg gccaacgtcc acacc

45

<210> 17

<211> 6

<212> DNA

<213> Unknown

<220>

<223> Synthetic Construct

<400> 17

gctagc

6

<210> 18

<211> 6

<212> DNA

<213> Unknown

<220>

<223> Synthetic Construct

<400> 18

aagctt

6

&lt;210&gt; 19

&lt;211&gt; 6

&lt;212&gt; DNA

&lt;213&gt; Unknown

&lt;220&gt;

&lt;223&gt; Synthetic Construct

&lt;400&gt; 19

tctaga

6

&lt;210&gt; 20

&lt;211&gt; 6

&lt;212&gt; DNA

&lt;213&gt; Unknown

&lt;220&gt;

&lt;223&gt; Synthetic Construct

&lt;400&gt; 20

tcgcga

6

&lt;210&gt; 21

&lt;211&gt; 21

&lt;212&gt; PRT

&lt;213&gt; Unknown

&lt;220&gt;

&lt;223&gt; Synthetic Construct

&lt;400&gt; 21

Asp	Tyr	Lys	Asp	Asp	Asp	Asp	Lys	Ala	Ser	Lys	Leu	Gly	Ala	Pro	Gly
1				5					10					15	

Pro	Leu	Gly	Ile	Ala
			20	

<210> 22

<211> 9

<212> PRT

<213> Unknown

<220>

<223> Synthetic Construct

<400> 22

Gly Pro Pro Gly Pro Pro Gly Leu Ala  
1 5

<210> 23

<211> 9

<212> PRT

<213> Unknown

<220>

<223> Synthetic Construct

<400> 23

Gly Pro Pro Gly Pro Pro Gly Glu Arg  
1 5

<210> 24

<211> 9

<212> PRT

<213> Unknown

<220>

<223> Synthetic Construct

<400> 24

Gly Pro Pro Gly Pro Pro Gly Pro Ala  
1 5

<210> 25

<211> 9



<212> PRT

<213> Unknown

<220>

<223> Synthetic Construct

<400> 25

Gly Pro Pro Gly Pro Pro Gly Ala Pro  
1 5

<210> 26

<211> 9

<212> PRT

<213> Unknown

<220>

<223> Collagen Type III

<400> 26

Gly Leu Ala Gly Ala Pro Gly Leu Arg  
1 5

<210> 27

<211> 45

<212> PRT

<213> Unknown

<220>

<223> Synthetic Construct

<400> 27

Gly Gly Lys Gly Asp Ala Gly Ala Pro Gly Glu Arg Gly Pro Pro Gly  
1 5 10 15

Leu Ala Gly Ala Pro Gly Leu Arg Gly Gly Ala Gly Pro Pro Gly Pro  
20 25 30

Glu Gly Gly Lys Gly Ala Ala Gly Pro Pro Gly Pro Pro  
35 40 45

<210> 28

<211> 7

<212> PRT

<213> Unknown

<220>

<223> Synthetic Construct

<400> 28

Met Gly Ala Pro Gly Ala Pro  
1 5

<210> 29

<211> 6

<212> PRT

<213> Unknown

<220>

<223> Synthetic Construct

<400> 29

Gly Ala Pro Gly Ala Pro  
1 5

<210> 30

<211> 7

<212> PRT

<213> Unknown

<220>

<223> Synthetic Construct

<400> 30

Gly Leu Glu Gly Pro Arg Gly  
1 5

<210> 31

Q67867 Sequence2.ST25.txt

<211> 26

<212> DNA

<213> Unknown

<220>

<223> Synthetic Construct

<400> 31

aattccatgg gtgctccagg tgctcc

26

<210> 32

<211> 22

<212> DNA

<213> Unknown

<220>

<223> Synthetic Construct

<400> 32

ggccacctgg tggacctggt gg

22

<210> 33

<211> 22

<212> DNA

<213> Unknown

<220>

<223> Synthetic Construct

<400> 33

ggcccgtggtg taagggtgac gc

22

<210> 34

<211> 20

<212> DNA

<213> Unknown

&lt;220&gt;

&lt;223&gt; Synthetic Construct

&lt;400&gt; 34

cgcgcacctg gtggacctgg

20

&lt;210&gt; 35

&lt;211&gt; 25

&lt;212&gt; DNA

&lt;213&gt; Unknown

&lt;220&gt;

&lt;223&gt; Synthetic Construct

&lt;400&gt; 35

cgcgcggttg taagggtgac gctgg

25

&lt;210&gt; 36

&lt;211&gt; 41

&lt;212&gt; DNA

&lt;213&gt; Unknown

&lt;220&gt;

&lt;223&gt; Synthetic Construct

&lt;400&gt; 36

acaaccctgg tggacctggt ggacctggtg gacctgggtg g

41

&lt;210&gt; 37

&lt;211&gt; 40

&lt;212&gt; DNA

&lt;213&gt; Unknown

&lt;220&gt;

&lt;223&gt; Synthetic Construct

&lt;400&gt; 37

ctagccccgc ggaccctcga gaccacaaca accctggtgg

40

&lt;210&gt; 38

&lt;211&gt; 9

&lt;212&gt; PRT

&lt;213&gt; Unknown

&lt;220&gt;

&lt;223&gt; Synthetic Construct

&lt;400&gt; 38

Gly Leu Ala Gly Ala Pro Gly Leu Arg  
1 5

&lt;210&gt; 39

&lt;211&gt; 1572

&lt;212&gt; DNA

&lt;213&gt; Unknown

&lt;220&gt;

&lt;223&gt; Collagen Type III Alpha I Chain

&lt;400&gt; 39

ccaggccac ttgggattgc tgggatcact ggagcacggg gtcttgcagg accaccaggc	60
atgccaggctc ctaggggaag ccctggccct caggggtgtca agggtgaaag tgggaaacca	120
ggagctaacg gtctcactgg agaacgtggc cccctggac cccagggtct tggtggtctg	180
gctggtacag ctggtgaacc tggaagagat ggaaaccctg gatcagatgg tcttccaggc	240
cgagatggat ctcttggtgg caagggtgat cgtggtgaaa atggctctcc tggtgcccct	300
ggcgctcctg gtcattccagg cccacctggc cctgtcgggc cagctggaaa gagtggtgac	360
agaggagaaa gtggccctgc tggccctgct ggtgctcccg gtcctgctgg tttccgaggt	420
gctcctgggc ctcaaggccc acgtggtgac aaagggtgaaa cagggtgaacg tggagctgct	480
ggcatcaaag gacatcgagg attccctggc aatccagggtg cccaggttc tccaggccct	540
gctggtcagc aggggtgcaat cggcagtgca ggacctgcag gcccagagg acctgttgga	600
cccagtgga ctcctggcaa agatggaacc agtggacatc cagggtcccat tggaccacca	660
gggcctcgag gtaacagagg tgaaagagga tctgagggtc cccagggcca cccagggcaa	720
ccaggccctc ctggacctcc tggtgcccct ggtccttgct gcgggtggtgt tggagccgct	780
gccattgctg ggattggagg tgaaaaagct ggcggttttg ccccgattta tggacctgaa	840
ccaatggatt tcaaaatcaa caccgatgag attatcactt cactcaagtc tggttaatgga	900

Q67867 Sequence2.ST25.txt

caaataaaaa	gcctcattag	tcctgatggg	tctcgtaaaa	accccgctag	aaactgcaga	960
gacctgaaat	tctgaccatc	ctgaactcaa	gactggagaa	tactgggtcg	accctaacca	1020
aggatgcaaa	ttggatgcta	tcaaggtatt	ctgtaatatg	gaaactgggg	aaacatgcat	1080
aagtgccaat	cctttgaatg	ttccacggaa	acactgggtg	acagattcta	gtgctgagaa	1140
gaaacacggt	tggtttggag	agtccatcga	tgggtggttt	cagtttagct	acggcaatcc	1200
tgaacttcct	gaagatgtcc	ttgatgtgca	gctggcattc	cctcgacttc	tctccagccg	1260
agcttcccag	aacatcacat	atcactgcaa	aaatagcatt	gcatacatgg	atcaggccag	1320
tggaaatgta	aagaaggccc	tgaagctgat	gggggtcaaat	gaagggtgaat	tcaaggctga	1380
aggaaatagc	aaattcacct	acacagttct	ggaggatggg	tgcacgaaac	acactgggaa	1440
tggagcaaaa	cagtctttga	atatcgaaca	cgcaaggctg	tgagactacc	tattgtagat	1500
attgcaccct	atgacattgg	tggtcctgat	caagaatttg	gtgtggacgt	tggccctggt	1560
tgctttttat	aa					1572

<210> 40

<211> 503

<212> DNA

<213> Unknown

<220>

<223> Synthetic Construct

<400> 40

gaattccatg	ggtgctccag	gtgctccagg	tggttaaggg	gacgctggg	ctccagggtga	60
aagagggtcca	ccagggtttg	ctggtgctcc	agggttgaga	ggtgggtgctg	gtccaccagg	120
tccagaagggt	ggtaaggggtg	ctgctgggtcc	accagggtcca	ccagggtgggc	ccgggtggtaa	180
gggtgacgct	ggtgctccag	gtgaaagagg	tccaccagggt	ttggctgggtg	ctccagggttt	240
gagagggtgg	gctgggtccac	cagggtccaga	agggtggaag	ggtgctgctg	gtccaccagg	300
tccaccagggt	gcgcgcgggtg	gtaaggggtga	cgctgggtgct	ccagggtgaaa	gagggtccacc	360
agggtttggct	ggtgctccag	gtttgagagg	tggtgctggg	ccaccagggtc	cagaagggtgg	420
taagggtgct	gctgggtccac	cagggtccacc	agggtccacca	ggtccaccag	gttggtgtgg	480
tctcgagggg	ccgcgggggct	agc				503

<210> 41

<211> 166

&lt;212&gt; PRT

&lt;213&gt; Unknown

&lt;220&gt;

&lt;223&gt; Synthetic Construct

&lt;400&gt; 41

Asn Ser Met Gly Ala Pro Gly Ala Pro Gly Gly Lys Gly Asp Ala Gly  
 1 5 10 15

Ala Pro Gly Glu Arg Gly Pro Pro Gly Leu Ala Gly Ala Pro Gly Leu  
 20 25 30

Arg Gly Gly Ala Gly Pro Pro Gly Pro Glu Gly Gly Lys Gly Ala Ala  
 35 40 45

Gly Pro Pro Gly Pro Pro Gly Gly Pro Gly Gly Lys Gly Asp Ala Gly  
 50 55 60

Ala Pro Gly Glu Arg Gly Pro Pro Gly Leu Ala Gly Ala Pro Gly Leu  
 65 70 75 80

Arg Gly Gly Ala Gly Pro Pro Gly Pro Glu Gly Gly Lys Gly Ala Ala  
 85 90 95

Gly Pro Pro Gly Pro Pro Gly Ala Arg Gly Gly Lys Gly Asp Ala Gly  
 100 105 110

Ala Pro Gly Glu Arg Gly Pro Pro Gly Leu Ala Gly Ala Pro Gly Leu  
 115 120 125

Arg Gly Gly Ala Gly Pro Pro Gly Pro Glu Gly Gly Lys Gly Ala Ala  
 130 135 140

Gly Pro Pro Gly Pro Pro Gly Pro Pro Gly Pro Gly Cys Cys Gly  
 145 150 155 160

Leu Glu Gly Pro Arg Gly  
 165

&lt;210&gt; 42

&lt;211&gt; 4

&lt;212&gt; PRT

<213> Unknown

<220>

<223> Endoplasmic Reticulum Retention Signal

<400> 42

Lys Asp Glu Leu  
1

<210> 43

<211> 4

<212> PRT

<213> Unknown

<220>

<223> Endoplasmic Reticulum Retention Signal

<400> 43

Lys Glu Glu Leu  
1

<210> 44

<211> 144

<212> DNA

<213> Unknown

<220>

<223> Synthetic Construct

<400> 44

agatccggtg gtaaggggtga cgctgggtgct ccaggtgaaa gaggtccacc aggtttggct 60

ggtgctccag gtttgagagg tgggtgctggt ccaccaggtc cagaaggtgg taaggggtgct 120

gctgggtccac caggtccacc aggt 144

<210> 45

<211> 144

<212> DNA

<213> Unknown



<220>

<223> Synthetic Construct

<400> 45

ccaccattcc cactgcgacc acgaggtcca ctttctccag gtggtccaaa ccgaccacga 60  
 ggtccaaact ctccaccacg accaggtggt ccaggtcttc caccattccc acgacgacca 120  
 ggtggtccag gtggtccatc tagg 144

<210> 46

<211> 48

<212> PRT

<213> Unknown

<220>

<223> Synthetic Construct

<400> 46

Arg Ser Gly Gly Lys Gly Asp Ala Gly Ala Pro Gly Glu Arg Gly Pro  
 1 5 10 15  
 Pro Gly Leu Ala Gly Ala Pro Gly Leu Arg Gly Gly Ala Gly Pro Pro  
 20 25 30  
 Gly Pro Glu Gly Gly Lys Gly Ala Ala Gly Pro Pro Gly Pro Pro Gly  
 35 40 45

<210> 47

<211> 8

<212> DNA

<213> Unknown

<220>

<223> Synthetic Construct

<400> 47

gatccggt

8

<210> 48

<211> 10

<212> DNA

<213> Unknown

<220>

<223> Synthetic Construct

<400> 48

gccatctagg

10

<210> 49

<211> 10

<212> DNA

<213> Unknown

<220>

<223> Synthetic Construct

<400> 49

agatccggta

10

<210> 50

<211> 9

<212> DNA

<213> Unknown

<220>

<223> Synthetic Construct

<400> 50

gccatctag

9

<210> 51

<211> 74

<212> PRT

<213> Unknown

<220>

Q67867 séquence2.ST25.txt

<223> Synthetic Construct

<400> 51

Asp Tyr Lys Asp Asp Asp Asp Asp Lys Glu Phe Leu Glu Pro Gly Arg  
1 5 10 15

Ser Gly Arg Ser Gly Gly Lys Gly Asp Ala Gly Ala Pro Gly Glu Arg  
20 25 30

Gly Pro Pro Gly Leu Ala Gly Ala Pro Gly Leu Arg Gly Gly Ala Gly  
35 40 45

Pro Pro Gly Pro Glu Gly Gly Lys Gly Ala Ala Gly Pro Pro Gly Pro  
50 55 60

Pro Gly Arg Ser Gly Pro Val Asp Pro Arg  
65 70

<210> 52

<211> 77

<212> PRT

<213> Unknown

<220>

<223> Synthetic Construct

<400> 52

Asp Tyr Lys Asp Asp Asp Asp Asp Lys Glu Phe Leu Glu Pro Arg Ser  
1 5 10 15

Gly Arg Ser Gly Gly Lys Gly Asp Ala Gly Ala Pro Gly Glu Arg Gly  
20 25 30

Pro Pro Gly Leu Ala Gly Ala Pro Gly Leu Arg Gly Gly Ala Gly Pro  
35 40 45

Pro Gly Pro Glu Gly Gly Lys Gly Ala Ala Gly Pro Pro Gly Pro Pro  
50 55 60

Gly Arg Ser Ile Asp Gly Ser Gly Pro Val Asp Pro Arg  
65 70 75